

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2005, 09:12:34 ; Search time 1914.49 Seconds
(without alignments)
1033.875 Million cell updates/sec

Title: US-10-798-192-4

Perfect score: 52
Sequence: 1 ggatgacacaaagacatg.....attactgtcttattatgc 52

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST.*
1: gb_est1.*
2: gb_est2.*
3: gb_est3.*
4: gb_est4.*
5: gb_est5.*
6: gb_est6.*
7: gb_est7.*
8: gb_est8.*
9: gb_est9.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	29.4	56.5	463	7	NS9414
C 2	29.2	56.2	601	9	CL376379
C 3	28.8	55.4	415	2	BE223173
C 4	28.8	53.8	323	5	BP745682
C 5	27.8	53.5	421	2	BE580731
C 6	27.8	53.5	742	8	AQ479841
C 7	27.6	53.1	563	2	BE240791
C 8	27.6	51.9	630	9	CR825712
C 9	26.8	51.5	212	7	CF588249
C 10	26.8	51.5	220	9	BX45244
C 11	26.8	51.5	471	8	BR421419
C 12	26.8	51.5	764	9	CMS05200
C 13	26.8	51.5	793	8	B2417682
C 14	26.8	51.5	814	8	B2612324
C 15	26.8	51.5	864	9	CG110241
C 16	26.8	51.5	1035	9	CW509683
C 17	26.8	51.5	1037	9	CL992336
C 18	26.8	51.5	1962	9	AG390939
C 19	26.6	51.2	544	8	B2112511
C 20	26.6	51.2	566	4	BJ528842
C 21	26.6	51.2	586	4	CL307851
C 22	26.6	51.2	600	7	CN747739
C 23	26.6	51.2	672	8	AZ062088
C 24	26.6	51.2	768	9	AG404371

NS9414	YF51e08.s1
CL376379	RPCL44.44
BE223173	kp73f05.y
BP745682	BP745682
BE580731	kp81a01.y
AQ479841	RPCL-11-2
BE240791	EST404840
CR825712	tigr-g88-
CF588249	USDA-FP.1
BX45244	Arabidops
BR421419	if80h05.b
AL137553	Tetradodon
IF80H05.G	
B2417682	
B2612324	WHA8V45TR
CG110241	PUBO407B
CW509683	ZMMBHC000
CL992336	ZMMBHC000
AG390939	Mus musci
B2112511	CH230-249
BJ528842	BJS28842
CL307851	02S0135-1
CN747739	SAL US008
AZ062088	RPF-23-4
AG404371	Mus musci

C 25	26.6	51.2	801	9	CC508846	CH240.351	CC508846	CH240.351
C 26	26.6	51.2	1005	7	CK282807	EST745529	CK282807	EST745529
C 27	26.4	50.8	629	8	AQ391541	CITBI-B1	AQ391541	CITBI-B1
C 28	26.4	50.8	794	8	BH201231	Sm1-54112	BH201231	Sm1-54112
C 29	26.4	50.8	856	7	CN168256	AGENCOURT	CN168256	AGENCOURT
C 30	26.2	50.4	398	7	CV350934	MK3-CT046	CV350934	MK3-CT046
C 31	26.2	50.4	406	7	CV350967	MK3-CT046	CV350967	MK3-CT046
C 32	26.2	50.4	562	8	AQ063007	CIT-HSP-2	AQ063007	CIT-HSP-2
C 33	26.2	50.4	644	9	CR330652	Medica90	CR330652	Medica90
C 34	26.2	50.4	684	9	CR330652	Medica90	CR330652	Medica90
C 35	26.2	50.4	809	8	B2380415	EBNCT0TR	B2380415	EBNCT0TR
C 36	26.2	50.4	820	8	AQ892006	HS_3100_B	AQ892006	HS_3100_B
C 37	26.2	50.4	1051	8	BH724762	BOMN175TF	BH724762	BOMN175TF
C 38	26.2	50.4	1051	8	BH724762	BOMN175TF	BH724762	BOMN175TF
C 39	26.2	50.4	200	5	BU710659	SJNAAC10	BU710659	SJNAAC10
C 40	26.2	50.4	255	6	CA993707	PAR09G03	CA993707	PAR09G03
C 41	26.2	50.4	464	1	AL384541	MEBC22G12	AL384541	MEBC22G12
C 42	26.2	50.4	464	1	AL384541	MEBC22G12	AL384541	MEBC22G12
C 43	26.2	50.4	581	2	BE239901	CE167365	BE239901	CE167365
C 44	26.2	50.4	637	9	CE598961	CE167365	CE598961	CE167365
C 45	26.2	50.4	674	6	CA965908	CE167365	CA965908	CE167365
C 46	26.2	50.4	689	9	CE734663	CE167365	CE734663	CE167365

ALIGNMENTS

RESULT 1
NS9414/C
LOCUS
DEFINITION
YV5108.81 Soares fetal liver spleen INFLS Homo sapiens CDNA clone
IMAGE:246278 3', mRNA sequence.

ACCESSION
NS9414.1
KEYWORDS
SOURCE
ORGANISM

REFERENCE
AUTHORS
Homo sapiens (human)
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homidae; Homo.

1 (bases 1 to 463)
Hillier, L., Lennon, G., Becker, M., Bonaudo, M.F., Chiappelli, B.,
Chissole, S., Dietrich, N., Dubuque, T., Favello, A., Glen, W.,
Hawkins, M., Hultman, M., Kucaba, T., Lacy, M., Le, N., Le, N.,
Mardis, E., Moore, B., Morris, M., Parsons, J., Prange, C., Rifkin, B.,
Rohtling, T., Schellenberg, K., Soares, M.B., Tan, F., Thierry-Mieg, J.,
Trevaekis, E., Underwood, K., Wohlmann, P., Waterston, R., Wilson, R.
and Maria, M.

Generation and analysis of 280,000 human expressed sequence tags
Genome Res. 6 (9), 807-828 (1996)
97044478

CONTACT: Wilson RK
Washington University School of Medicine
444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810

Email: est@wustl.wustl.edu
This clone is available royalty-free through LBL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert length: 1494 Std Error: 0.00
Seq primer: m13 -40 forward
High quality sequence stop: 444.

FEATURES
Location/Qualifiers
1..463
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:3795524"
/db_xref="taxon:9606"
/clone="IMAGE:246278"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares fetal liver spleen INFLS"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2005, 02:55:23 ; Search time 239,707 Seconds
(without alignments)
1284.177 Million cell updates/sec

Title: US-10-798-192-4

Perfect score: 52
Sequence: 1 ggtgacacacaaagacatg.....attactgtctcttattgac 52

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4390206 seqs, 2959870667 residues

Total number of hits satisfying chosen parameters: 8780412

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Database :
1: Genesegq16dec04:*
2: genesegq19808:*
3: genesegq19908:*
4: genesegq2000a:*
5: genesegq2001a:*
6: genesegq2002a:*
7: genesegq2002b:*
8: genesegq2003a:*
9: genesegq2003b:*
10: genesegq2003c:*
11: genesegq2003d:*
12: genesegq2004a:*
13: genesegq2004b:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	44	84.6	307	10	ADP42898
2	44	84.6	4404	10	ADL13987
3	44	84.6	4642	12	ADG38762
4	44	84.6	4652	6	AAZ49209
5	44	84.6	4652	6	AA517712
6	31	59.6	2307	3	AAZ49212
7	27.4	52.7	5041	3	AAH18245
8	26.6	51.2	2477	4	AAH17555
9	26.4	50.8	208700	13	ABD32688
10	26.4	50.8	243390	13	ABD33366
11	25.8	49.6	404	4	AA191464
12	25.8	49.6	435	10	ADP02845
13	25.8	49.6	2736	11	AD131648
14	25.8	49.6	3090	2	AAQ85925
15	25.8	49.6	3492	6	AB167678
16	25.8	49.6	3492	8	AAU52788
17	25.8	49.6	80321	11	ACN45012
18	25.6	49.2	1732	4	AB125255
19	25.6	49.2	3732	4	AB125254
20	25.6	49.2	135827	13	ABD33219

C	21	25.4	48.8	756	10	ACC61741	Acc61741 Gene sequ
C	22	25.4	48.8	756	10	ADK64037	ADK64037 Disease t
C	23	25.2	48.5	3364	4	AB112736	Ab112736 Drosophi
C	24	25.2	48.5	32572	6	AA517820	AA517820 PAdV-5 HN
C	25	25.2	48.5	47243	11	ACN44690	ACN44690 Human gen
C	26	25.2	48.5	177851	8	AA157272	AA157272 bA438B23-
C	27	25	48.1	766	4	AA195559	AA195559 Human neu
C	28	25	48.1	6179	4	AA546343	AA546343 Tumour su
C	29	25	48.1	6179	6	ABK31250	ABK31250 Signal tr
C	30	25	48.1	18491	4	AAK65236	AAK65236 Human imm
C	31	25	48.1	18491	8	ABZ73979	ABZ73979 Secreted
C	32	25	48.1	18491	10	ABR16919	ABR16919 Human sec
C	33	25	48.1	18491	10	ABZ7560	ABZ7560 Human sec
C	34	25	48.1	18495	4	AAK65235	AAK65235 Human imm
C	35	25	48.1	18495	8	ABZ73980	ABZ73980 Secreted
C	36	25	48.1	18495	10	ABR16920	ABR16920 Human sec
C	37	25	48.1	18495	10	ABZ67561	ABZ67561 Human sec
C	38	25	48.1	25806	4	AAK67566	AAK67566 Human imm
C	39	25	48.1	40783	12	ADQ97173	ADQ97173 Human can
C	40	25	48.1	107330	12	ADQ97316	ADQ97316 Mouse can
C	41	25	48.1	116592	8	ABX15519	ABX15519 Human tyr
C	42	25	48.1	116592	10	AD47900	AD47900 Human tita
C	43	25	48.1	335913	5	AA161371	AA161371 Soybean 2
C	44	25	48.1	335913	5	AA161372	AA161372 Soybean 2
C	45	24.8	47.7	11520	4	AB141118	AB141118 Drosophi

ALIGNMENTS

RESULT 1
ADP42898/c
ID ADP42898 standard; DNA; 307 BP.
XX
AC ADP42898;
XX
DT 12-FEB-2004 (first entry)
XX
DE AAV5 cap protein derived intron DNA.
XX
DE helper construct; packaging; Cap protein; serotype A; serotype B; intron;
KM P40 promoter; P5 promoter; P19 promoter; Rep protein; gene therapy; de.
XX
OS Adeno-associated virus 5.
XX
PM WO2003074686-A1.
XX
PD 12-SEP-2003.
XX
PF 07-MAR-2003; 2003WO-EP002351.
XX
PT 07-MAR-2002; 2002DE-01010139.
XX
PR (MEDI-) MEDIGENE AG.
XX
PA Hoerer M, Dubielzig R, Ries S, Kosfeld-Bergauer M;
XX
PI WPI, 2003-690032/65.
XX
DR New helper construct for packaging adeno-associated virus, useful for
XX preparing viral particles for gene therapy, encodes Cap and Rep genes of
XX different serotypes.
XX
PS Disclosure: Fig 10; 43pp; German.
XX
XX This invention describes a novel helper construct for packaging
XX recombinant adeno-associated virus (AAV) in which the sequence encoding
XX the Cap protein of serotype A is linked to regulatory regions of a
XX serotype B AAV, the Cap protein intron is of serotype A and the sequences
XX encoding Cap and Rep are functionally separate. The invention also
XX describes a system for packaging rAAV comprising a helper construct that
XX encodes a functional Rep from a serotype other than A, particularly B and
XX a host cell for packaging rAAV of serotype A containing a copy of the

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2005, 09:12:34 ; Search time 1914.49 seconds
(without alignments)
1033.875 Million cell updates/sec

Title: US-10-798-192-3

Sequence: 1 cgcataaagaacagcaat.....catgtcttctgtcacc 52

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 34239544 seqs, 19032134700 residues

Total number of hits satisfying chosen parameters: 68479088

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database :
EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_nuc:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gest1:*
9: gb_gest2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	29.2	56.2	601	9	CL376379 RPT44.44
2	28.8	55.4	415	2	BE223173
3	28.4	54.6	463	7	NS9414
4	27.8	53.5	421	2	BE580731
5	27.8	53.5	742	8	AO479841
6	27.6	53.1	323	5	BP745682
7	27.6	53.1	563	2	BE240791
8	27.6	53.1	630	9	CE825712
9	26.8	51.5	420	9	BE945244
10	26.8	51.5	471	8	BZ421419
11	26.8	51.5	764	8	CN805200
12	26.8	51.5	793	8	BZ417682
13	26.8	51.5	814	8	BZ612324
14	26.8	51.5	864	9	CG110241
15	26.8	51.5	1035	9	CW509683
16	26.8	51.5	1037	9	CL992336
17	26.8	51.5	1962	4	AG390999
18	26.6	51.2	566	4	BJS28842
19	26.6	51.2	600	7	CN747739
20	26.6	51.2	1005	7	CK282807
21	26.4	50.8	212	7	CF588249
22	26.4	50.8	794	8	BH201231
23	26.4	50.8	856	7	CN168256
24	26.4	50.8	888	5	BU410024

25	26.4	50.8	991	5	BU224056
26	26.2	50.4	338	7	CV350934
27	26.2	50.4	406	7	CV350967
28	26.2	50.4	544	8	BZ212511
29	26.2	50.4	586	9	CL307851
30	26.2	50.4	672	8	AZ062088
31	26.2	50.4	809	8	BZ390415
32	26.2	50.4	1051	8	BH724762
33	26.2	50.4	200	5	BU710659
34	26.2	50.4	255	6	CA993707
35	26.2	50.4	464	1	AL384541
36	26.2	50.4	497	9	CE157365
37	26.2	50.4	581	2	BE239901
38	26.2	50.4	637	9	CE598961
39	26.2	50.4	642	9	CE546561
40	26.2	50.4	701	2	AW774794
41	26.2	50.4	794	5	BO164967
42	26.2	50.4	797	5	BU770122
43	26.2	50.4	868	8	BH159163
44	26.2	50.4	874	8	BH150358
45	26.2	50.4	885	8	AZ550717

ALIGNMENTS

RESULT 1
LOCUS CL376379/c
DEFINITION RPT44.444113.r RPT44 Sus scrofa genomic clone RPT44.444113,
genomic survey sequence.
ACCESSION CL376379
VERSION CL376379.1 GI:51428344
KEYWORDS GSS.
SOURCE Sus scrofa (pig)
ORGANISM Sus scrofa
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Cetartiodactyla; Suidae; Sus.
1 (bases 1 to 601)

REFERENCE Rogatcheva, M.B., Meyers, S., He, W., Larkin, D.M., Marron, B.M.,
Beaver, J.B., and Schook, L.B.
Piggy-BACing the Human Genome: Constructing a Porcine Physical Map
Through Comparative Genomics
Unpublished (2004)
Other GSSs: RPT44.444113.f
Contact: Lawrence B. Schook
Department of Animal Sciences
University of Illinois at Urbana Champaign
1201 W. Gregory Dr., Urbana, IL 61801, USA
Tel: 217 265 5326
Fax: 217 244 5617
Email: schook@uiuc.edu

Clones are derived from the porcine BAC library RPT44
(http://www.bacpac.chori.org/porcine242.htm). For BAC library
availability, please contact Pieter de Jong (pje@chori.org).
Clones may be purchased from BACPAC Resources
(http://BACPACresources.chori.org). This work was undertaken as part
of the International Swine Genome Sequencing Consortium by
University of Illinois at Urbana Champaign, USA with funds provided
by grant No. AG2002-34480-11828 from USDA-CSREES and
AG2001-35205-09965 from USDA/NRI (Livestock Genome Sequencing
Initiative)

Class: BAC ends.
Seq primer: SP6
Plate: 444 row: I column: 13

FEATURES
source Location/Qualifiers

1..601
/organism="Sus scrofa"
/mol_type="genomic DNA"
/strain="four pigs (Breed: 37.5% Yorkshire Landrace and 25%
Weishan)"
/db_xref="taxon:9823"
/clone="RPT44.444113"

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 18, 2005, 06:10:24; Search time 506.39 Seconds
(without alignment)
2679.250 Million cell updates/sec

Title: US-10-798-192-1

Perfect score: 28

Sequence: 1 gagcaataatgatttaaccagtgatg 28

Scoring table: IDENTITY NUC
Gapop 10.0, Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues 9416466

Total number of hits satisfying chosen parameters:

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank:
1: gb_ba:*
2: gb_hg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	26.4	94.3	2370	6	AX753285 Sequence
2	26.4	94.3	2489	6	AX753280 Sequence
3	26.4	94.3	2495	6	AX753281 Sequence
4	26.4	94.3	2501	6	AX753275 Sequence
5	26.4	94.3	2504	6	AX753278 Sequence
6	26.4	94.3	2969	6	AX753303 Sequence
7	26.4	94.3	3084	6	AX753284 Sequence
8	26.4	94.3	3095	6	AX753286 Sequence
9	26.4	94.3	3098	6	AX753254 Sequence
10	26.4	94.3	3098	6	AX753255 Sequence
11	26.4	94.3	3098	6	AX753282 Sequence
12	26.4	94.3	3102	6	AX753262 Sequence
13	26.4	94.3	3105	6	AX753264 Sequence
14	26.4	94.3	3105	6	AX753265 Sequence
15	26.4	94.3	3106	6	AX753266 Sequence
16	26.4	94.3	3106	6	AX753261 Sequence
17	26.4	94.3	3106	6	AX753263 Sequence
18	26.4	94.3	3106	6	AX753279 Sequence
19	26.4	94.3	3113	6	AX753276 Sequence

20	26.4	94.3	3113	6	AX753277 Sequence
21	26.4	94.3	3117	6	AX753288 Sequence
22	26.4	94.3	3121	6	AX753289 Sequence
23	26.4	94.3	3122	6	AX753287 Sequence
24	26.4	94.3	3122	6	AX753290 Sequence
25	26.4	94.3	3123	6	AX753286 Sequence
26	26.4	94.3	3128	6	AX753272 Sequence
27	26.4	94.3	3128	6	AX753273 Sequence
28	26.4	94.3	3128	6	AX753291 Sequence
29	26.4	94.3	3128	6	AX753292 Sequence
30	26.4	94.3	3129	6	AX753304 Sequence
31	26.4	94.3	3131	6	AX753259 Sequence
32	26.4	94.3	3197	6	AX753274 Sequence
33	26.4	94.3	3276	6	AX753283 Sequence
34	26.4	94.3	4385	6	AX753250 Sequence
35	26.4	94.3	4393	6	AX753249 Sequence
36	26.4	94.3	4393	14	AF513852 Adeno-889
37	26.4	94.3	4683	6	BD242775 Adeno-889
38	26.4	94.3	4683	6	AR562507 Sequence
39	26.4	94.3	4683	14	AF028704 Adeno-889
40	26.4	94.3	4718	6	BD242766 Adeno-889
41	26.4	94.3	4718	6	AR562498 Sequence
42	26.4	94.3	4718	6	AX753251 Sequence
43	26.4	94.3	4718	14	AF063497 Adeno-889
44	26.4	94.3	4721	6	AX753246 Sequence
45	26.4	94.3	4721	14	AF513851 Adeno-889

ALIGNMENTS

RESULT 1
LOCUS AX753285 2370 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 40 from Patent EP1310571.
ACCESSION AX753285
VERSION AX753285.1 GI:32166142
KEYWORDS
SOURCE unidentified
ORGANISM unidentified
REFERENCE 1
AUTHORS Gao, G., Wilson, J.M. and Alvira, M.
TITLE A method of detecting and/or identifying adeno-associated virus (AAV) sequences and isolating novel sequences identified thereby
JOURNAL Patent: EP 1310571-A 40 14-May-2003;
The Trustees of The University of Pennsylvania (US)
FEATURES
source 1..2370
/organism="unidentified"
/mol_type="unassigned DNA"
/db_xref="taxon:32644"
/note="new AAV serotype, clone 43.5"
ORIGIN
Query Match 94.3%; Score 26.4; DB 6; Length 2370;
Best Local Similarity 96.4%; Pred. No. 6;
Matches 27; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 GAGCAATAATGATTAAACGAGTATG 28
DB 819 GAGCAATAATGACTTAAACGAGTATG 846
RESULT 2
LOCUS AX753280 2489 bp DNA linear PAT 23-JUN-2003
DEFINITION Sequence 35 from Patent EP1310571.
ACCESSION AX753280
VERSION AX753280.1 GI:32166137
KEYWORDS
SOURCE unidentified
ORGANISM unidentified

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using SW model

Run on: September 18, 2005, 06:10:24 ; Search time 578.732 Seconds
(without alignments)
2679.250 Million cell updates/sec

Title: US-10-798-192-2

Perfect score: 32

Sequence: 1 gcttagaccgcatgcacgaagctttatcgc 32

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4708233 seqs, 24227607955 residues

Total number of hits satisfying chosen parameters: 9416466

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenEmbl:*

1: gb_da:*

2: gb_hcg:*

3: gb_in:*

4: gb_om:*

5: gb_ov:*

6: gb_pac:*

7: gb_ph:*

8: gb_pl:*

9: gb_dr:*

10: gb_ro:*

11: gb_rst:*

12: gb_sy:*

13: gb_un:*

14: gb_vi:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	25	78.1	4718	6	BD242766 Adeno-ss
2	25	78.1	4718	6	AR562498 Sequence
3	25	78.1	4718	6	AX753251 Sequence
4	25	78.1	4718	14	AF063497 Adeno-ss
5	22	68.8	62421	8	NCB384
6	21	65.6	205656	2	AL55531 Neurospor
7	20.8	65.0	223802	2	AC130671 Mus muscu
8	20.4	63.7	106778	5	AL592062 Zebrafish
9	20.4	63.7	110000	2	Continuation (3 of
10	20.4	63.7	165002	2	AC150642 Bos tauru
11	20.4	63.7	177531	2	AC129120 Rattus no
12	20.4	63.7	233510	2	AC136569 Rattus no
13	20.4	63.7	238007	2	AC097745 Rattus no
14	19.8	61.9	145302	10	AL663080 Mouse DNA
15	19.8	61.9	148546	9	AC087256 Homo sapi
16	19.8	61.9	168227	9	AC021868 Homo sapi
17	19.8	61.9	173326	2	AC131775 Mus muscu
18	19.8	61.9	193437	2	AC119524 Rattus no
19	19.8	61.9	196817	9	AC019209 Homo sapi

C 20	19.8	61.9	218928	2	AC107228 Mus muscu
C 21	19.8	61.9	225858	2	AC127878 Rattus no
C 22	19.8	61.9	234160	2	AC103270 Rattus no
C 23	19.8	61.9	240965	2	AC126054 Mus muscu
C 24	19.8	61.9	253048	2	AC128497 Rattus no
C 25	19.8	61.9	349688	1	CR378671 Photobact
C 26	19.6	61.3	199350	10	AL772381 Mouse DNA
C 27	19.6	61.3	223449	2	AC12864 Rattus no
C 28	19.4	60.6	38042	9	HSV461C10 Human DNA s
C 29	19.4	60.6	49433	9	AL157707 Human DNA
C 30	19.4	60.6	62372	2	AC091054 Homo sapi
C 31	19.4	60.6	92148	8	AP006089 Lotus cor
C 32	19.4	60.6	105468	8	AP004852 Oryza sat
C 33	19.4	60.6	122103	9	HS291J10 Human DNA s
C 34	19.4	60.6	134258	8	AC136227 Oryza sat
C 35	19.4	60.6	136607	9	AC083809 Homo sapi
C 36	19.4	60.6	153088	9	AC097654 Homo sapi
C 37	19.4	60.6	16052	8	AC136224 Oryza sat
C 38	19.4	60.6	168968	2	AC010161 Homo sapi
C 39	19.4	60.6	169323	2	AL161906 Homo sapi
C 40	19.4	60.6	171430	2	AC027160 Homo sapi
C 41	19.4	60.6	172228	10	AC114671 Mus muscu
C 42	19.4	60.6	206512	2	AC015544 Homo sapi
C 43	19.4	60.6	219331	2	AC134216 Rattus no
C 44	19.4	60.6	219910	2	AC141018 Rattus no
C 45	19.4	60.6	225644	2	AC102598 Mus muscu

ALIGNMENTS

RESULT 1	BD242766/c	4718 bp	DNA	linear	PAT 17-JUL-2003
LOCUS	BD242766				
DEFINITION	Adeno-associated virus serum type 1 nucleic acid sequence, vector				
ACCESSION	BD242766				
VERSION	BD242766.1	GI:33052536			
KEYWORDS	JP 2002529098-A/1.				
SOURCE	unidentified				
ORGANISM	unclassified.				
REFERENCE	1 (bases 1 to 4718)				
AUTHORS	Wilson,J.M. and Xiao,W.				
TITLE	Adeno-associated virus serum type 1 nucleic acid sequence, vector				
JOURNAL	Patent: JP 2002529098-A 1 10-SEP-2002;				
COMMENT	THE TRUSTEES OF THE UNIVERSITY OF PENNSYLVANIA				
OS	AAV-1				
PN	JP 2002529098-A/1				
PD	10-SEP-2002				
PR	02-NOV-1999 JP 2000581227				
PI	05-NOV-1998 US 60/107114				
PC	JAMES M WILSON,WEIDONG XIAO				
PC	C12N15/09,A61K31/711,A61K48/00,A61P43/00,C12N1/15,C12N1/19,PC				
CC	Adeno-associated virus serum type 1 nucleic acid sequence, CC				
FT	vector and host				
FT	cell containing the same				
FT	Location/Qualifiers				
FT	Key				
FT	CDS				
FT	Location/Qualifiers				
FEATURES	1..4718				
SOURCE	/organism="unidentified"				
ORIGIN	/mol_type="genomic DNA"				
Query Match	78.1%; Score 25; DB 6; Length 4718;				
Best Local Similarity	100.0%; Pred. No. 0.51;				
Matches	25; Conservative 0; Mismatches 0; Indels 0; Gaps 0;				